ERIC Document Resume | See Instruction Sheet | DATE OF RESUME 6/4/65 | DATE RECEIVED

A. PUBLICATION DATA

1. ERIC NUMBER
ED 000 689

2. ORIGINATOR CONTROL NO(S).
8133

3. IDENTIFICATION NO.

4. AUTHOR(S):
Rising, Edward J.

5. TITLE OF DOCUMENT ABSTRACTED:
The Use of Minimum Cost Motion Pictures for Engineering Instruction

6. PUBLICATION TITLE:
N.A.

7. EDITOR(S):
N.A.

8. PLACE OF ORIGIN
ORG. Univ. of Iowa
ST. CITY, STATE Iowa City, Iowa
TELE. EXT.
INST/LOC CODE

9. PUBLISHER OR SOURCE
University of Iowa

10. VOL.	11. PAGES	12. MO./YR.	13. TOTAL PP.
N.A.	N.A.	Aug. 59	79

DO NOT WRITE IN SPACE BELOW

MICROFILM COPY AVAILABLE
YES ☒ NO ☐

COPYRIGHT GRANTED
YES ☐ NO ☒

COPYRIGHT RELEASE FILED
YES ☐ NO ☐

TYPE OF REPORT
PROPOSAL ☐
IN-PROGRESS ☐
FINAL ☒
FOLLOW-UP ☐

OTHER ERIC REFERENCES:

ERIC SATELLITE CODE

B. INDEXING DATA

14. SUBJECT AREA

15. TYPE OF ENDEAVOR
Doctoral dissertation and experimental study

16. INDEXING TERMS (6-12)

experimental study	engineering edu.	economic factors
higher edu.	teaching methods	effectiveness
professional edu.	audio-visual aids	

17. ABSTRACT (MAX. 250 WORDS)

The purpose of the experiment was to evaluate the usefulness of minimum-cost motion picture films as instructional tools for teaching a computation laboratory course in engineering. The project consisted of three controlled experiments which were run over a 3-week study unit in which all the lectures of the project sections were given by minimum-cost films.

As a result of the experiment, the author concluded that this type of film offers promise of being economically feasible on a self-amortizing basis for repetitive, multiple section courses, helpful for teacher training, and useful for extending the influence of superior teachers.

ERIC No.

18. SOURCE OF ABSTRACT

C. FUNDING DATA

19. SUPPORTING ORGANIZATION

20. DIVISION OR DEPT.

21. CONTRACT NO.

PROJECT NO.

22. TOTAL COST OF PROJECT

23. DURATION OR PROJECT

U.S. Department of Health, Education, and Welfare
Office of Education

ERIC Form 1

R-6- R-6-18 (Engineering)

8133

THE PRODUCTION AND USE OF MOTION PICTURES FOR ENGINEERING INSTRUCTION*

by

Edward James Rising

SUMMARY

So that the usefulness of "minimum cost" motion picture films as instructional tools for teaching a computation laboratory course in Engineering might be evaluated, three controlled experiments were run over a three-week study unit in which all the lectures of the treatment sections were given by minimum cost films. The findings were interpreted to indicate that there was no difference between the performance on criterion tests of matched treatment and control sections. This conclusion was reached on the basis of accepting the null difference hypothesis; the greatest weakness in the investigation was that the available sample size was rather small for this statistical technique. Two criterion tests were given, one immediately following the study unit, measuring the effect of learning, one two months later, measuring the retention. The minimum estimate of the reliability of any criterion test was .69.

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

INTRODUCTION

If the educational process in multiple-section courses is thought of as a man-machine system, typically the machine portion is rudimentary (chalk and blackboard) and the man portion is frequently tedious and repetitive. If the machine-role can be expanded, the repetition and tedium can be reduced in the man-role. Motion pictures provide a method of doing this and many other things. A completely filmed course of lectures will provide the advantages of broadening the influence of the more talented teachers, of easing the shortage of trained teachers, of reducing the per student instructional cost, and creating a basis for teacher training within professional areas.

---

*This is a portion of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Industrial and Management Engineering in the Graduate College of the State University of Iowa, August 1959.

ED000689

This investigation, covering a study unit which takes three weeks to complete in the course Engineering Problems as taught at the State University of Iowa, is a pilot study of the efficacy of using minimum cost motion picture films to carry the major instructional burden throughout a whole course. The course in which the experimentation was performed is required of all freshmen students and is given two semester hours credit. It meets twice a week during a sixteen-week semester for three hours each meeting as a computation laboratory. The material covered in this course is the use of logarithms and the slide rule to perform computation processes that are basic to engineering and the physical sciences. These techniques are applied to problems of conversions of units from one system to another, exponential equations of various types, and the solution of triangles and other problems of a similar nature.

The psychological reasons for choosing this course for experimentation were that the students taking the course have a wide range of ability levels because it is given to freshmen students, and the sample size for the statistical populations was generous because it is a required course, and, finally, the course is of a mathematical nature, and, consequently, representative of a type of course common in the engineering curriculum. The practical reasons for selecting this course were that it is a multiple section course, the content is stable; it is an expensive course in terms of teacher and student time; and, consequently, any advantage we obtain, educational or financial, will produce more total benefit.

The method of producing film for this investigation was suggested by the work of Greenhill and Kepler. Their concept was adapted to fit the needs of higher education. The films of this study were produced by the trained camera crews using the extensive facilities and multi-camera, synchronous-sound techniques of the State University of Iowa film production unit. The presentations that were filmed were essentially classroom lectures reorganized for the medium and rehearsed for smooth presentation.

STATISTICAL OUTLINE

The technique most often employed for the statistical design of methods experiments is to formulate a straw man in the form of a null hypothesis, e.g., "There is no difference between the treatments as measured by a criterion test." The "desirable" outcome of the experiment is to reject the null difference hypothesis and conclude that the treatments must be unequal, and proceed to detail the inequality. In back of this technique is the notion that a proposed (new) method is of little interest if it is not possible to reject the null difference hypothesis, i.e., if the new method is not demonstratably superior to the old method in current use.

The logic of statistical inference supports this approach, desirable to reject non-reject abandon, because of the nature of the two types of error to which every statistical test is subjected, i.e., Type I, rejecting a hypothesis when it is true, and Type II, accepting a hypothesis when it is false. If an experimental treatment is abandoned when the null difference hypothesis cannot be rejected, the practical consequence of a Type II (acceptance) error has been eliminated. Only taking action on the basis of rejected hypotheses eliminates any practical risk from an acceptance (Type II) error. Type I (rejection) error offers no problem because its risk is set in advance as the decision rule.

This investigation is unlike most methods experiments in that it is more vulnerable to Type II errors because a "desirable" outcome is acceptance of the null hypothesis. It is argued that administrative advantages, including decreased costs, make it desirable to change to the new film teaching method if it can be done without loss of quality. Of the three possible outcomes of a comparison between the old and the new methods, (a) the old method is demonstratably better than the new method, (b) the two methods are statistically indistinguishable, (c) the new method is demonstratably better than the old method, it then follows that either of the latter two conclusions are "desirable" from the point of view of the new method.

EXPERIMENTAL METHOD

Six films were produced to cover completely the content material of the three-week instructional unit on logarithms which is the first instructional unit in the course. Data were collected during both semesters of the 1958-59 school year; the student enrollment was the random sample of a population defined as the students, past, present, and future who take the course.

The engineering ability of the students was obtained, for purposes of the experiment, from the Pre-Engineering Ability Test (Form ZPA) which was administered the first day classes met so that the measure of engineering aptitude dated at the start of the experiment. The treatment and control groups were matched by the counting-off method (one for me, one for you, etc. starting from the top of the list of students arranged by engineering aptitude), and the ability levels within the treatments were constituted by dividing the matched groups into the upper, middle, and lower third based on engineering aptitude. The criterion measure for the experiments was a unit examination, given immediately following the study unit, and a portion of the final examination, to measure retention. Estimates of the reliabilities of these tests vary between .69 and .80. The parameters of these tests are described in Table I.

Experiment I, during the Fall semester, was an attempt to find out whether the films could provide a presentation as useful to the students as similar lectures given live to a matched control group by an experienced teacher; there was also an attempt to determine whether any improvement could be gained by a second showing of the films. The treatment groups in this experiment were designated I and II. The Treatment I group viewed the filmed lectures; the Treatment II group viewed the filmed lectures twice to test the effect of repeated showing; both of these treatments were compared to a matched control group.

Experiment II, during the Spring semester, was to compare the results obtained from a treatment group conducted by a carefully supervised undergraduate teaching assistant, using the films and duplicated lecture notes, to the results obtained from the control group conducted by an experienced teacher presenting similar lectures live.

The control treatment in both these experiments consisted of live lectures given at the same time as the film treatments by the same person that appeared on the film. Choosing the same person who appeared on the films to make the live lectures for the control group was an attempt to hold down the variability due to different teachers. In this way, the test comes nearer to being a test of the usefulness of the films compared to live presentations.

Experiment III, also during the Spring semester, was an attempt to determine whether the films and notes could be used in a class conducted by an experienced teacher (called the treatment group) to produce the same effect as another experienced senior staff member using his own methods of teaching (called the control group). This is the only experiment in which the control group was not conducted by the teacher appearing on the films, and, therefore, helps to "ground" the experiment. The former experiments may show differences exist, but this experiment indicates that the level at which we are making these comparisons is reasonably high.

It seemed desirable to make several types of comparisons with the data; first, to compare the average performance of the experimental and control groups; second, if a difference is shown to exist in average performance, to determine whether the difference is general throughout the ability range of the students or selective in a particular ability range;

third, if no difference exists between the average scores, to examine the data to see whether this may be caused by one ability level scoring disproportionately higher on one treatment and another ability level scoring disproportionately higher in other treatments so that the average remains the same.

DISCUSSION

The results of Experiment I, given in Table II and III, may be summarized as follows: It appears that (a) a generally experienced instructor never having taught the course before can achieve the same average results using minimum cost films as can an experienced instructor, teaching in person, who has taught the course many times; (b) there is no apparent benefit from showing the films a second time; (c) the results are similar for the unit examination, which measures the effect of learning, and the final examination, which measures retention.

The same "average" results can, of course, be achieved many ways; for example, the high ability students of one treatment may score disproportionately higher than the high-ability students of other treatments which might be balanced by the low-ability students of the latter treatments scoring higher than the low students of the former treatment. This would result in a selectivity between the various treatments and the ability levels which appears in the analysis as interaction. No interaction was found in the unit examination data, and it is argued that the interaction observed in final examination data is not an "in fact characteristic" of the parent population.

The results of the tests made with the data from the final examinations for the presence of interaction can be accounted for because of the large amount of data "lost" from the experiment due to students withdrawing from the course. The pattern of this "lost" data shows many more students dropping from the lower ability levels. Since the data suggests that even within the lowest ability category it is the poorer students that are dropping, this lowest ability category has a systematic upward bias, (i.e., elimination of the lowest individuals from a category raises the average score of the group.) This bias should account for the interaction effect that was detected.

The numerical results of Experiment II were much the same as Experiment I. The interpretation of the results of Experiment II, given in Table II and III, may be summarized as follows: (a) a carefully supervised undergraduate teaching assistant, using films and duplicated lecture notes, can produce the same effect as an experienced teacher giving the same lecture in person, (b) this result is similar for the unit examination, which measures the effect of learning, and the final examination, which measures retention.

The final examination data of this experiment do not present a clear picture because approximately 15% of this data have been "lost."

Two impressions were obtained during the course of this experiment. First, little effort, in addition to carefully previewing the film with the teaching assistant, is necessary to properly instruct the teaching assistant to conduct any particular class. Second, the teaching assistant appeared competent to introduce and summarize film presentations. These impressions suggest that it would be possible for one experienced teacher to utilize and control a number of teaching assistants. The fact that the duplicated lecture notes, used during this experiment, were accepted and used by the students in the same manner as personally taken notes indicates that the difficulties which students usually have taking notes from film can be satisfactorily overcome with material that can be prepared in advance.

The results of Experiment III, given in Table II and III, may be summarized as follows: (a) Two experienced teachers both having taught the course many times before, one using the films and duplicated notes and the other his own teaching methods, can achieve the same results; (b) these results are similar for the unit examination, which measures the effect of learning and the final examination, which measures retention.

In this experiment there are only four cases in each ability level category for each treatment (cell) to start with, and there is a very high percentage of data "lost" before the final examination, therefore, the results of these calculations should not be considered conclusive. It is interesting to note that the results, for what they are worth, are compatible with experiment I and II in that there is no significant difference between treatments and similar to Experiment I in that there is observed interaction between treatments and levels. In the Experiment III final examination data, however, there is no significant difference between levels which is probably due to the systematic upward bias on the lower ability level cells caused by the poorer students dropping the course after the unit examination and before the final examination.

COSTS

The expendable supplies required for the production of the six films used in this study cost an average of $260.53 per film. A careful estimate of the total cost of labor to produce each film, including camera men, film editor, and processing techinicians but not including the teacher-actor, was made at $100.00 per film. Subsequent to this experiment, fifteen additional films on slide rule computations were produced. Expendable supplies for these films averaged $290.00 per film and an estimate of $100.00 per film for technical labor, as before, seems appropriate for these later films.

During the academic year, teaching assistants, under the immediate supervision of a senior staff member who will indicentally be conducting a teacher development program with these assistants, will use these films in combination with duplicated lecture notes to present all the course material to the students. It is estimated the $1750.00 will be saved during one academic year. Thus, the cost of producing the films can be quickly amortized even if the computation does not include any estimate for income that may be obtained from the sale of these films.

CONCLUSIONS

Two major conclusions seem to emerge from the evidence that has been discussed: First, students can learn both conceptual and factual material from this type of film, and they apparently learn as well as from the same lecture presented live under similar circumstances; and, second, this type of film offers a promise of being (a) economically feasible on a self-amortizing basis for repetitive, multiple section courses, (b) useful for extending the influence of superior teachers, and (c) helpful for teacher training.

BIBLIOGRAPHY

1. Rising, E. J., "The Use of Minimum Cost Motion Pictures for Engineering Instruction.," Unpublished Dissertation, State University of Iowa, Iowa City, Iowa, 1959.

2. Greenhill, L. P., Kepler, L. F., "A Study of the Feasibility of Local Production of Minimum Cost Sound Motion Pictures, *SPECDEVCEN* Project 20-E-4, 1955.

3. Kumata, Kideya, *An Inventory of Instructional Television Research*, Educational Television and Radio Center, Ann Arbor, Michigan, 1956.

4. Dale, E., Dunn, F. W., Hoban, C. F. Jr., Schneider, E. *Motion Pictures in Education*. American Council on Education Committee on Motion Pictures in Education, H. W. Wilson, 1937.

5. Hoban, C. F., Jr., Van Ormer, E. B., "Instructional Film Research," *Technical Report No. SDC 269-7-19*, 1951

6. Lindquist, E. F., *Design and Analysis of Experiments in Psychology and Education*, Houghton Mifflin Company, 1953.

TABLE I Description of the examinations used as criterion measures at the end of the study unit and as final examinations for Experiments I, II, and III.

		Experiment I		Experiment II & III	
CHARACTERISTIC	Ideals	Unit Test	Final	Unit Test	Final
DISCRIMINATIONS					
High (.40 and Up)	more than 25%	34%	50%	2.5%	14%
Moderate (.20 to .39)	more than 25%	40%	45%	37.5%	36%
Low (.01 to .19)	less than 15%	14%	5%	57.5%	41%
Zero or Negative	less than 5%	12%	0%	2.5%	9%
SCORE STATISTICS					
Mean		28.79	15.28	29.13	16.71
	(indicated)[a]	(25.78)	(13.75)	(25.22)	(13.75)
Standard Deviation		5.13	3.76	5.23	2.99
	(indicated)[b]	(5.08)	(2.75)	(4.92)	(2.75)
Reliability	more than .70	.784	.798	.803	.690
Probable Error	no standard	1.60	1.14	1.57	1.12

a - Midpoint of range between highest possible and expected chance score.

b - One-sixth of range between highest possible and expected chance score.

TABLE II Values of the mean square ratios for "Treatments," "Levels," and "Interaction" (T x L) for Experiments I, II, and III.

	Mean Square Ratios					
	d. f.	Treatments / Within	d. f.	Levels / Within	d. f.	T x L / Within
Experiment I						
unit exam	$\frac{2}{87}$	2.20	$\frac{2}{87}$	17.19*	$\frac{4}{87}$	0.72
final exam	$\frac{2}{82}$	1.03	$\frac{2}{82}$	13.80*	$\frac{4}{82}$	7.42*
Experiment II						
unit exam	$\frac{1}{33}$	.00	$\frac{2}{33}$	14.66*	$\frac{2}{33}$	.06
final exam	$\frac{1}{30}$	.46	$\frac{2}{30}$	.42	$\frac{2}{30}$	1.03
Experiment III						
unit exam	$\frac{1}{18}$	1.11	$\frac{2}{18}$	8.38*	$\frac{2}{18}$	.98
final exam	$\frac{1}{13}$	.84	$\frac{2}{13}$	1.74	$\frac{2}{13}$	5.56*

* Greater than the corresponding value of "F" at 5% probability.

TABLE III The mean value of criterion measure for the different treatment and levels for the three experiments.

		Unit Exam Data				Final Exam Data			
		Control	Treatment I	Treatment II	Level Average	Control	Treatment I	Treatment II	Level Average
Experiment I	High	33.79	30.93	31.80	32.27	19.00	16.79	17.40	17.82
	Medium	28.36	28.64	26.00	28.12	14.14	16.79	9.60	14.58
	Low	27.21	24.14	25.20	25.61	14.00	13.00	16.00	13.88
	Treatment Average	29.79	27.90	27.67		15.71	15.52	14.33	
Experiment II	High	32.71	33.14		32.93	17.57	18.43		18.00
	Medium	30.86	30.43		30.64	18.00	16.43		17.21
	Low	26.29	26.14		26.21	18.00	17.29		17.64
	Treatment Average	29.95	29.90			17.86	17.38		
Experiment III	High	30.50	29.25		29.88	18.75	15.75		17.25
	Medium	29.25	31.50		30.38	13.25	18.25		15.75
	Low	21.75	25.50		23.62	17.00	11.25		14.12
	Treatment Average	27.17	28.75			16.33	15.08		

Sample Size Note: Experiment I began with 14 cases in each cell of the Treatment I and control groups and 5 cases per cell in the Treatment II group. Experiment II started with 7 cases per cell and Experiment III with 4 cases per cell in all groups. Because several students dropped the course before completion some data was "lost" to the analysis.

#B
R-6

THE USE OF MINIMUM COST MOTION PICTURES FOR ENGINEERING INSTRUCTION

by

Edward James Rising

An Abstract

Of a dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Department of Industrial and Management
Engineering in the Graduate College of
the State University of Iowa

August, 1959

Chairman: Professor Wellborn R. Hudson

ABSTRACT

So that the usefulness of "minimum cost" motion picture films as instructional tools for teaching a computation laboratory course in Engineering might be evaluated, three controlled experiments were run over a three week study unit in which all the lectures of the treatment sections were given by minimum cost films. The findings were interpreted to indicate that there was no difference between the performance on criterion tests of matched treatment and control sections. This conclusion was reached on the basis of accepting the null difference hypothesis; the greatest weakness in the investigation was that the available sample size was rather small for this statistical technique. Two criterion tests were given, one immediately following the study unit, measuring the effect of learning, one two months later, measuring the retention. The minimum estimate of the reliability of any criterion test was .69. Estimates are developed for time and cost and compared for the present and proposed method. Non-economic advantages of the proposed method are discussed.

THE USE OF MINIMUM COST MOTION PICTURES
FOR ENGINEERING INSTRUCTION

by

Edward James Rising

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Department of Industrial and Management
Engineering in the Graduate College of
the State University of Iowa

August, 1959

Chairman: Professor Wellborn R. Hudson

ACKNOWLEDGMENTS

Grateful acknowledgement is due to many people who made this study possible: Dean Emeritus Dawson for his encouragement and support, Professor Deegan for his inspiration and encouragement in the early phases of this investigation, Professor Howe for his administrative support, Professor Cochran and Mr. Nelson of the Audio Visual section of the Extension Division for cooperation and help producing the necessary films, Dr. Mittman of the University Testing Service for help with the examinations, and Professor Schneider, Mr. Warnock and Mr. Mittleberg for their cooperation in the classroom to make the study possible.

Special mention should be made of Professor Hudson who was advisor for this investigation. He knew when to help and when to force independent work. This is no small thing.

TABLE OF CONTENTS

page
SUMMARY 1
PURPOSE 2
INTRODUCTION 3
  Background 3
  Job Analysis 5
  Economic Aspects 8
  Non-Economic Aspects 14
METHOD 17
  Summary of Method 26
INSTRUMENTATION 31
RESULTS 34
DISCUSSION 45
  Experiment I 45
  Experiment II 51
  Experiment III 55
CONCLUSIONS 58
FUTURE RESEARCH 59
APPENDIX A (The Course) 61
APPENDIX B (Statistical Design Considerations) 62
APPENDIX C (Computations for Economic Analysis) 71
APPENDIX D (Film Production) 75
BIBLIOGRAPHY 78

TABLE OF TABLES

Number

I A comparison of the costs of the Present and the Proposed Methods of teaching five sections of Engineering Problems for one year.................. 10

II A comparison of the average costs of the Present and the Proposed Methods of teaching various numbers of sections of Engineering Problems for one year.................................. 10

III Descriptions of the examinations used as measuring instruments at the end of the study unit and as final examinations for Experiments I, II, and III......................... 33

IV The Results of Experiment I......................... 36

V The Results of Experiment II......................... 39

VI The Results of Experiment III......................... 42

VII Comparison of the "best estimates" of the range and mean of the Pre-Engineering Aptitude scores of the Parent and Hypothetical Populations of the various experiments............................ 66

VIII Comparison of Pre-Engineering Aptitude raw scores for the various experimental sections for the Fall Semester Sample............................ 69

IX Comparison of Sample to Norm Groups 44a

X Coorelation coeficients between Pre-Engineering Aptitude Test and the Unit Test and Final Examination raw scores 44b

TABLE OF FIGURES

Number

page

1 Diagram of the treatment by levels design used for Experiment I, during the Fall semester portion of the investigation, showing the size of the sample drawn for each cell 22

2 Diagram showing the meeting time of the various treatment groups for Experiment I, during the fall semester 22

3 Diagram of the treatment by levels design used for Experiment II, during the Spring semester portion of the investigation, showing the size of the sample for each cell.................................. 25

4 Diagram of the treatment by levels design used for Experiment III, during the Spring semester portion of the investigation, showing the size of the sample for each cell 25

5 Set Diagram 77

SUMMARY

So that the usefulness of "minimum cost" motion picture films as instructional tools for teaching a computation laboratory course in Engineering might be evaluated, three controlled experiments were run over a three week study unit in which all the lectures of the treatment sections were given by minimum cost films. The findings were interpreted to indicate that there was no difference between the performance on criterion tests of matched treatment and control sections. This conclusion was reached on the basis of accepting the null difference hypothesis; the greatest weakness in the investigation was that the available sample size was rather small for this statistical technique. Two criterion tests were given, one immediately following the study unit, measuring the effect of learning, one two months later, measuring the retention. The minimum estimate of the reliability of any criterion test was .69. Estimates are developed for time and cost and compared for the present and proposed method. Non-economic advantages of the proposed method are discussed.

PURPOSE

High schools, junior colleges and small liberal arts colleges prepare many of their students for professional curricula, usually at other institutions. If whole courses of lectures were available on film, produced by the professional schools, the students in the pre-professional schools could be given an introduction to professional school discipline earlier in their academic career. These films would demonstrate to the faculties of the pre-professional institutions, more vividly than is now possible, the type of work that will be required of their students when they enter the professional schools.

In addition to this new use for instructional film, and its advantages, a completely filmed course of lectures provides the additional advantages of broadening the influence of the more talented teachers, of providing a basis for teacher training within professional areas, of easing the shortage of trained teachers, of reducing the per student instructional cost, and of providing a means to utilize more fully the existing classroom space.

This investigation, covering a three week study unit in Engineering Problems (59:2), is a pilot study of the efficacy of using minimum cost motion picture films to carry the major instructional burden throughout a whole course.

INTRODUCTION

Background

Motion pictures, as an educational tool, have been the subject of research for about forty years; several excellent summaries of this work are available (1,2,3).* It is generally conceded that well-made educational films facilitate both factual and conceptual learning (2,3), and are indeed superior in this regard to a poor instructor, equal to an average instructor, and perhaps equal to the better teachers; and films accomplish their result in a shorter time (2,3).

Because television research has usually been done with longer instructional units it may, in some ways, be more applicable to those seeking information about how sets of films can be of help with the teaching of whole courses. However, television research itself is usually not as precise as motion picture research because at this stage of development of this medium research into learning is often an afterthought to an ambitious production problem. Much of the serious research done with television is directed toward

*The numbers in parenthesis, for example, (1,2,3) refer to the serially numbered works cited in the bibliography found on page 78. The bibliography is arranged in the order that the works are first cited.

problems of communication between audience and teacher (4); this two-way communication introduces a variable that limits the generalizations that can be applied to motion pictures (5).

A survey of pertinent literature indicates that instructional film researchers have avoided dealing with whole courses in preference to a hodge-podge of instructional units. This may be partially explained: perhaps the researchers feel motion pictures have virtue as teaching aids; and, in an effort to prove this, they select only material well adapted to the medium. Perhaps funds available to one project, or researcher, were not adequate to produce enough films to cover a whole course. The evident lack of research at the college level, particularly in the area of physical science and engineering, may be because the material is sufficiently eruidite that content area specialists must supervise the work. Up to now this task has been the responsibility of medium specialists. Whatever the reason for the situation, the research at the college level devoted to physical science or engineering is fragmentary and unsystematic (6).

The idea for this investigation was adapted from the work of Greenhill and Kepler (7), who studied and described the feasability of locally producing films with untrained camera crews using a do-it-yourself camera kit. This concept was adapted for this investigation to fit the needs of higher education. The films were produced by the trained camera crews using the extensive facilities of the university. The presentations that were filmed were essentially classroom lectures slightly reorganized for the medium and rehearsed for smooth presentation.

Job Analysis

If the process of teaching multiple section courses is analyzed in accordance with Management Engineering principles, specific possibilities for improving effectiveness of faculty utilization are suggested.

If films were available for the whole course, it would be possible to re-allocate the work of instructing so that the professional level work is done by professional level people, and the routine work is done by assistants. In addition to savings of money and professional time, which will subsequently be shown to be promising, such a division will enhance the usefulness of the professional people and increase their satisfaction with their work because it relieves them of spending a large part of their time doing routine work. The division of work also provides the opportunity to develop orderly training for the assistants which will ensure competent instruction for the students. Through a system of teaching assistantships, this system will provide a means of helping the best of the advanced undergraduates or beginning graduate students financially and at the same time encouraging them to enter the teaching profession.

With these things in mind, it is helpful to separate the work of teaching the course into the two following categories.

Professional Level Work

1. Train assistants
2. Supervise construction and administration of exams.
3. Plan
4. Organize the content and manner of presentation of content material
5. Assign grades
6. Act as representative of Department, College, and University to the student.

Work of the Assistants

1. Operate projection equipment.
2. Distribute and correct assignments.
3. Answer questions about assignments.

The present method may be compared to the proposed method in terms of these functions.

Present Method

In the present method an experienced teacher for each section uses few, if any, teaching aids. All the functions, both professional and sub-professional, are carried out by this individual. The teacher must have, at a minimum, the ability to do the operation calling for the highest level of skill listed.

Proposed Method

In the proposed method one or more functional groups, each consisting of an experienced teacher, a teaching assistant for each of several sections, and a set of filmed presentations is a "teaching team." It would be the responsibility of the experienced teacher to do the planning, to train the teaching assistants, to supervise closely the construction and administration of all examinations, to assign all grades, and to be responsible for revising old films and producing new films in a planned manner. The teaching assistants would meet the individual sections, introduce and follow up the films, distribute and correct assignments, answer questions about the problems, and help individuals as time permits. To maintain "control" of the process, the supervising teacher would meet regularly with the teaching assistants as a group and individually with them as necessary; he would visit classes to support and aid the assistant with his day-to-day problems and be generally available to help with the personal and academic problems of the students who, being freshmen, may require informal contact with the

University. This type of administration is analogous to the "exception principle" of administration which is widely practiced in industry.

Under the proposed method, the definition of the roles of the supervising teacher and the assistants is intended to place professional work with professional persons and the routine work with the assistants. Broadly speaking, the supervising teacher is responsible for the planning and supervision functions which, for example, might be identified as a) handling exceptions to routine, b) initiating new routine, c) handling problems beyond the competence of assistants, d) monitoring the program for incipient malfunction, and e) guiding the professional development of the assistants. The assistant is generally charged with implementation of the program which, for example, might be identified functionally as a) performing tasks as directed, b) reporting exceptions to the supervising teacher, c) reporting problems beyond his competence to the supervising teacher, d) showing interest and effort toward professional development.

Economic Aspects

The economics of money and professional time of the present and proposed methods can best be examined from an example. The following comparison is based on the current situation at the State University of Iowa. Although each application will need to be examined separately, these figures are typical of computation laboratories which have yearly enrollments of about 200 divided into two semesters.

Table I shows a cost comparison of the present vs. the proposed method for five sections for one year. These figures show a saving of $1500 to $2000 annually, which is a saving of from $300 to $400 per section each year. Table II summarizes the comparison for four, five, and six sections for one year. The figures for the present method in Table II are developed from the average section costs taken from Table V ($1200 per section); the average cost for one teaching assistant ($450 per section) is used to develop the information for the proposed method. This table shows a savings of $750 dollars yearly for each additional section over five per year. The computations are made on this basis because very little additional work will be required from the supervising teacher for one or two additional sections once the job is organized for the first few sections, and certainly the additional sections would not require an increase in the funds budgeted for remaking films and filming new topics.

TABLE I A comparison of the costs of the Present and the Proposed Methods of teaching five sections of Engineering Problems for one year. (These figures are developed in Appendix C.)

	Present Method	Proposed Method
Permanent Staff Salaries	$ 6000	$ 1600
Teaching Assistant Salaries	none	2000 to 2500
Cost of remaking films and making films on new topics	none	400
	$ 6000	$4000 to 4500

The saving per year from the Proposed Method is $1500 to $2000.

TABLE II A comparison of the average costs of the Present and the Proposed Methods of teaching various numbers of sections of Engineering Problems for one year.

Number of Sections	Present Method Yearly Cost	Proposed Method Yearly Cost	Yearly Saving	Amortization Time (Years)
4	$ 4800	$ 3800	$ 1000	5+
5	6000	4250	1750	3±
6	7200	4700	2500	2+
7	8400	5150	3250	2-

The Amortization Time is computed on the basis of "departmental out of pocket costs", and, the films being used only at the institution at which they were produced.

It is known that costs will change with enrollment; however, the comparison between the present and the proposed method in Table II shows that the proposed method indicates increased savings with an increased enrollment. Although these estimates are presented with confidence, no data are currently available that will describe cost-benefit relationships with any precision, or how these relationships will change with enrollment. As the estimates are extrapolated, the conclusions become less dependable.

The cost of production of the films is an important item in these computations. If a department of the State University of Iowa had made the six films for this study, the billing would have averaged $260.53 each. The remainder of Engineering Problems could be covered by about fourteen to sixteen more films of comparable costs; the total billing for the production of films for the whole course would then be in the order of $5750. At present enrollment figures, this amount could be amortized in just over three years. If enrollments increase as everyone expects, the amortization time would be shorter.

These figures, however, do not reflect the true cost of making the films. The true cost should include the time of the lecturer to reorganize his material for the medium and rehearse his lectures. The salaries of the production crew and film editor, which are currently paid from the University budget, and some reasonable figure for overhead should be included. All of this may double the above cost figures. It should be pointed out, however, that from the point of view of a departmental administrator, the $5750 is an "out-of-pocket" cost, but the increases noted are not "out-of-pocket" costs. It is probable that sets of films could be rented or sold to other institutions which would partially offset, or perhaps even cover completely, the costs of producing these films.

On the basis of the figures available, it is impossible to make definite statements about amortization time of total costs. As yet we do not have any knowledge of the expected sales or rental revenue, or enough experience producing this type of film to make precise estimates of total cost. It does appear, however, that "out-of-pocket" costs could be amortized in something just over three years even if the film is used only within this University.

This method of analysis, which demonstrates that dollar savings are possible from the proposed method, can also be used to demonstrate that time of professional people can be saved. When five sections of this course are run for a full year by the present method, it is estimated that 1280 hours of work are required from professional level people (See Appendix C for details of these time estimates). It is estimated that the proposed method will require 320 hours of work from a professional level person for the same five sections for one year. The estimated saving of the proposed method over the present method, in terms of professional time only, is 960 hours per year at the present enrollment level.

It is estimated that the planning, rehearsal, shooting, and review necessary to produce a single film takes about 15 hours for the teacher-actor, and about 5 hours for another teacher who acts as critic. On this basis, for 22 films covering the whole course, the total investment of professional time to produce the necessary films is in the order of 440 hours.

With an expenditure of 440 hours of professional time, it appears to be possible to save approximately 960 hours of professional level time in one academic year. The amortization time is seen to be (440/960 = .46) less than one-half year, which can be considered less than one semester. The investment of professional level effort and time required to make these films can be seen as an investment that is quickly returned.

Non-Economic Aspects

The advantages that are available from a full course of lectures and an organization of the teaching duties as previously described are not all economic.

The influence of a teacher can be extended with this type of films to more students than possible without them. The best teachers, if their lectures can be permanently recorded on film, can be presented to many students in many institutions. The lectures that the students receive from films of this sort will be better, on the average, than this same teacher could give in a classroom because he can afford extra time for preparation and rehearsal if he knows a lecture is to be preserved. This extra effort takes time that cannot be justified for a one time presentation to a small class.

Late registering students and students who are absent from class can get the same lecture as their classmates without an additional drain on the time of the teacher. It is also possible for students who have difficulty with certain topics to review certain lectures at a later time at their convenience with little or no inconvenience to the teacher.

Using films in place of lectures provides a classroom situation that can be utilized for training inexperienced teachers with little or no loss of teaching effectiveness because the students are always assured of uniformly high quality lectures. Beginning graduate students and the best of the advanced undergraduate students, intending to enter graduate school, who plan to teach to support themselves while working for advanced degrees can be given careful guidance, help, and training with teaching techniques and problems. Because many departments plan to have several graduate students engaged in teaching, often in lower level courses, an increase in their teaching skill will represent a beneficial improvement in the teaching of the less mature and more numerous underclassmen. If these young teachers have successful and rewarding experiences teaching, they will be more interested in teaching as a career.

The assembled data indicate teaching with film and assistants will produce the same measured outcome as can be obtained with an experienced teacher alone. These results, because they were obtained from the very first set of films produced by this technique and the first assistant that was trained to use these films, may not do justice to the teaching value of this type of film. It is entirely possible that, when more experience is gained in making and showing this type of film and in selecting and training assistants to use the films and to administer sections, the effectiveness of the instruction of the proposed method of film teaching will improve so that it may be on the average superior to traditional methods.

The organization and rehearsal of material for a film, together with the opportunity to view the result, is a real help to any teacher who is interested in improving his own classroom lectures. Use of the resulting film in a classroom provides an opportunity to see whether the ideas and techniques that the lecturer tried to communicate were really assimilated by the class. The questions from students, or lack of questions, the performance of the students, or their lack of performance, can be very helpful to a teacher with a sincere desire to improve his technique. To know that a certain point was covered on a film and to have questions about that point from several students is a commentary on the effectiveness of a portion of the presentation. The lack of questions about a point not specifically covered may imply that most of the students are able to reason out from general principles that are presented. It can be seen, then, that because making and using of films can be helpful to the senior staff members who desire to improve their lecturing techniques, the teacher training aspect of the proposed film teaching method need not be limited to beginning level teachers.

METHOD

To implement the primary aim of this research, which is to collect evidence about the instructional efficacy of minimum cost motion pictures, the investigation was carried out in three phases: a comparison between the immediate performance of students taught by minimum cost films and those taught by usual instructional techniques, a comparison between the retention of the material by the students taught by this type of film and those taught by the usual method, and an analysis of the circumstances under which this type of film is economical to produce and use as an educational aid.

Six films were produced to cover completely the content material of a three week instructional unit on logarithms for Engineering Problems (59:2) as presently taught at the State University of Iowa. (See Appendix A.) Data were collected during both semesters of the 1958-59 school year.

The technique most often employed for the statistical design of methods experiments is to formulate a straw man in the form of a null hypothesis, e.g. "There is no difference between the treatments as measured by a criterion test". The "desirable" outcome of the experiment is to reject the null difference hypothesis and conclude that the treatments must be unequal, and procede to detail the inequality. In back of this technique is the notion that a proposed (new) method is of little interest if it is not possible to reject the null difference hypothesis, i.e., if the new method is not demonstratably superior to the old method in current use.

The logic of statistical inference supports this, "desirable to reject, non-reject abandon", approach because of the nature of the two types of error to which every statistical test is vulnerable, i.e., Type I, rejecting a hypothesis when it is true, and Type II, accepting a hypothesis when it is false. If there is no more interest in the experiment when the null hypothesis has not been rejected, the practical consequences of a Type II (acceptance) error has been eliminated. Being only interested in action on the basis of rejected hypotheses eliminates any practical risk from the type of error inherent in an acceptance decision. Type I (rejection) error, of course, offers no problem because its risk is set in advance as the decision rule.

This investigation is unlike most methods experiments in that it is more vulnerable to Type II errors because a "desirable" outcome is acceptance of the null hypothesis.

It will be argued that administrative advantages make it desirable to change to the new film teaching method if it can be done without loss of quality. It then follows, that of the three possible outcomes of a comparison between the old and the new method, a) that the old method is demonstratably better than the new method, b) the two methods are statistically indistinguishable, c) that the new method is demonstratably better than the old method, either of the latter two conclusions are "desirable" from the point of view of the new method.

It needs to be clearly understood that when an null hypothesis is accepted, one is accepting something that is fundamentally an assumption that cannot be disproved by the available data. No probability can be set on the Type II error involved in a decision to accept an hypothesis. This risk is present because, for the distributions compared, the "within" variability may be large enough to conceal a "between" variability (if it exists) for the sample that is drawn. The weakest point of the experimental work presented here is the small sample size that was available to perform these comparisons.

In order to perform a statistical study of films as a teaching device the specific hypotheses developed must be relative to a numerical measure of either the result of learning or the learning itself. Because the most important goal of Engineering Problems is to develop the ability of the student to compute the solutions of certain types of problems quickly and accurately, a timed examination on these problems is an inherently valid measure of the most important outcome of the learning, if not a measure of the learning itself. Such an examination was given at the end of the three week instructional unit, and again approximately two months later during the final examination. These two examinations allow comparisons to be made for immediate learning, from the first examination, and for retention, from the final examination.

It seems desirable to make several types of comparisons with the data; first, to compare the average performance of the experimental and control groups; second, if a difference is shown to exist in average performance, to determine whether the difference is general throughout the ability range of the students or selective in a particular ability range; third, if no difference exists between the average scores, to examine the data to see whether this may be caused by one ability level scoring high on one treatment and another ability level scoring high in other treatments. All these tests can be accomplished with the "treatment by levels" design; an added advantage is that the assumptions required by this design are less restrictive than most other designs. A detailed consideration of the statistical design, and the assumptions involved in this application, is given in Appendix B.

Ability levels were obtained for purposes of the experiment from the Pre-Engineering Ability Test (Form ZPA). The test was administered the first day the classes met, so the measure of engineering aptitude dated at the start of the experiment; on the basis of these scores the students were assigned to one of the three ability levels.

Three separate experiments were attempted during the course of the study. Because of the uniformity that occured in the distribution of the ability levels from section to section during the fall semester, it was decided to pool all the data from the sections into one experiment. (For a detailed discussion of the assumptions and how they are met, see Appendix B.) Two different treatments, called Film I and Film II, and the Control treatments were combined into this experiment; it is referred to as Experiment I and is diagrammed in Figure 1.

The distribution of ability levels was unbalanced throughout the sections during the spring semester; more of the high ability students appeared in the morning sections. It was possible to balance the two morning sections by the

		Control	Film I	Film II
Engineering Ability Levels	High	n = 14	n = 14	n = 5
	Medium	n = 14	n = 14	n = 5
	Low	n = 14	n = 14	n = 5

Figure 1. Diagram of the treatment by levels design used for Experiment I, during the Fall semester portion of the investigation, showing the size of the sample drawn for each cell.

	Mon.	Tues.	Wed.	Thurs.	Fri.
7:30		Film I and Control		Film I and Control	
8:30					
9:30					
10:30					
11:30					
1:30					
2:30	Film II	Film I and Control		Film I and Control	Film II
3:30					
4:30					
5:30					

Figure 2. Diagram showing the meeting time of the various treatment groups for Experiment I, during the fall semester.

counting-off method, one section for treatment and one section for control, and to balance the two afternoon sections in the same manner, but because of scheduling difficulties it was impossible to obtain balance between the morning and afternoon groups. Consequently, two separate experiments were performed with the data taken during the spring semester. The one performed with the morning groups is referred to as Experiment II, and the one in the afternoon as Experiment III. The single section meeting on Monday and Friday could not be used for experimentation because it could not be matched with a control group except by methods of covariance. Covariance analysis is considered undesirable for the purposes of this study because it does not permit examination of interaction and involves restrictive assumptions.

Experiment I, during the first semester, was an attempt to find out whether the films could provide a presentation as useful to the students as lectures from an experienced teacher, and to attempt to determine whether any improvement could be gained by a second showing of the films. Showing the film instead of the usual lecture is referred to as the Film I treatment, and showing the film instead of the usual lecture and reshowing it later in the period after the students had worked for a short time on their daily problem is referred to as the Film II treatment. The Control treatment consisted of live lectures given at the same time as the Film I treatment by the same person who appeared on the film. Figure 1 shows a diagram of this experiment giving the sample size in each cell, and Figure 2 shows the hours at which the various treatment groups met. The Control and Film I treatment were given to the pairs of sections meeting at the same hour, and both morning and afternoon sections were included; the Film II treatment was given to a single section meeting on different days.

Experiment II, during the spring semester, attempted to find out whether a carefully supervised undergraduate teaching assistant, using the films and duplicated lecture notes, could handle a section of students as successfully as an experienced teacher giving, live, the same lecture as is on the film. The same "teacher" who appears on the film teaches the control group; duplicated lecture notes are not used. Figure 3 shows a diagram of this design giving the sample size for each cell.

Experiment III, also during the spring semester, is an attempt to determine whether the films and notes, used as in Experiment II, produce the same effect as an experienced senior staff member using his own methods of teaching.

		Treatment Categories	
		Control	Film
Engineering Ability Levels	High	n = 7	n = 7
	Medium	n = 7	n = 7
	Low	n = 7	n = 7

Figure 3. Diagram of the treatment by levels design used for Experiment II, during the Spring Semester portion of the investigation, showing the size of the sample for each cell.

		Treatment Categories	
		Control	Film
Engineering Ability Levels	High	n = 4	n = 4
	Medium	n = 4	n = 4
	Low	n = 4	n = 4

Figure 4. Diagram of the treatment by levels design used for Experiment III, during the Spring Semester portion of the investigation, showing the size of the sample for each cell.

This experiment is the only one in which the non-film group is not taught by the same person who appears on the films, and in which the non-film group is not given live lectures that duplicate the film presentation. This experiment is diagramed in Figure 4.

Summary of Method

The following outline summarizes the hypotheses considered in this investigation. It should be noted that they are written in a "null difference" form and the subsequent experimental work will demonstrate that it is impossible to reject them with the present sample data. The technique of showing treatment equivalence by failure to reject null difference hypotheses is not a strong technique with small sample sizes. The available sample size is the outstanding weakness of this investigation.

I. Experiment I is based on the following treatment groups and hypotheses.

   A. Treatment Groups

      1. The Film I treatment consisted of two sections of students who viewed a motion picture instead of a live lecture. These sections were conducted by a generally experienced Instructor who had never taught this course before.

2. The Film II treatment consisted of one section of students who viewed a motion picture instead of a live lecture and, after working a while on their daily problem, viewed the film a second time. This section was conducted by an experienced Instructor who had taught the course before, and, further, this Instructor was the one who acted as the "teacher" on the films.
3. The Control treatment consisted of two sections of students who were given a live lecture that duplicated the film presentation. These sections were conducted by the same person who was the "teacher" on the films.

B. Hypotheses

1. There is no difference between the criterion measures of the various treatment groups:
    a. When the criterion measure is the result of the examination measuring the immediate effects of learning;
    b. When the criterion measure is the result of the final examination used to measure retention.
2. There is no interaction between the criterion measure for the various treatment groups and the ability levels:

        a. When the criterion measure is the results of the examination measuring the immediate effects of learning;
        b. When the criterion measure is the results of the final examination used to measure retention.

II. Experiment II is based on the following treatment groups and hypotheses.

  A. Treatment Groups.
    1. The treatment consisted of one section of students who viewed the films instead of a live lecture. This section was conducted by an undergraduate Teaching Assistant who was given careful supervision.
    2. The control treatment consisted of one section of students who were given a live lecture that duplicated the film presentation. This section was conducted by the same person who appeared as "teacher" on the films.

  B. Hypotheses
    1. There is no difference between the criterion measures of the various treatment groups:
        a. When the criterion measure is the result of the examination measuring the effects of learning;

    b. When the criterion measure is the result of the final examination used to measure retention.
  2. There is no interaction between the criterion measure for the various treatment groups and the ability levels:
    a. When the criterion measure is the result of the examination measuring the effects of learning;
    b. When the criterion measure is the result of the final examination used to measure retention.

III. Experiment III is based on the following treatment groups and hypotheses.

A. Treatment Groups

1. The film treatment consisted of one section of students who viewed the films instead of the live lecture. This section was conducted by the same person who appeared on the films as the "teacher".
2. The control treatment consisted of one section of students who were taught by an experienced teacher, other than the person who appeared on the films, using his own methods. This teacher

was an Associate Professor who had taught this course many times before.

B. Hypotheses

1. There is no difference between the criterion measures of the treatment groups:
    a. When the criterion measure is the result of the examination measuring the effects of learning;
    b. When the criterion measure is the result of the final examination used to measure retention.
2. There is no interaction between the criterion measure for the various treatment groups and the ability levels:
    a. When the criterion measure is the result of the examination measuring the effect of learning;
    b. When the criterion measure is the result of the final examination used to measure retention.

INSTRUMENTATION

Three different measuring instruments were used during this investigation, a test to stratify the sample into ability levels, a criterion test to measure effect of learning, and a final examination to measure retention. The measure used to stratify the sample into the three ability levels was the Pre-Engineering Ability Test (Form ZPA); that is, "ability" was defined for purposes of the experiment in terms of a score on this test. This test has been in use for a number of years. It is published by the Educational Testing Service, and it was developed from the old Pre-Engineering Inventory, a six hour test which originated as a joint project of the Engineering Council for Professional Development and the American Society for Engineering Education. The test in its present form is an eighty minute multiple choice test given in two separately timed parts; the first part deals with the comprehension of scientific materials, and the second part deals with general mathematical ability. It was designed for use by college freshmen and high school graduates. The test is reported to have a reliability of .90 as estimated by the Kuder-Richardson Formula 20. The standard error of measurement is estimated as 3.9.

The criterion tests, used to measure the effect of the learning and retention, were first constructed and used during the term preceding this investigation. Both tests were then improved on the bases of item analyses. The unit examination is a 1 hour multiple choice test of about 40 items; the final examination is a 35 minute multiple choice test of about 25 items. The estimates of the parameters of these tests are given in Table III. The estimates of the reliabilities of all the tests are all above .69. The means and the standard deviations of the scores on all the tests were near the ideal values, which indicates that reasonable score distributions were obtained on all tests.

TABLE III Description of the examinations used as measuring instruments at the end of the study unit and as final examinations for Experiments I, II, and III.

Test Numbers		580921	590202	590310	590603
CHARACTERISTIC	Ideals	Unit Test	Final	Unit Test	Final
RELEVANCE					
Content Details	0%	0%	0%	0%	0%
Vocabulary	less than 20%	0%	0%	0%	0%
Facts	less than 20%	0%	0%	0%	0%
Generalizations	more than 10%	0%	0%	0%	0%
Explanations	more than 10%	0%	0%	0%	0%
Applications	more than 10%	100%	100%	100%	100%
DISCRIMINATIONS					
High (.40 and up)	more than 25%	34%	50%	2.5%	14%
Moderate (.20 to .39)	more than 25%	40%	45%	37.5%	36%
Low (.01 to .19)	less than 15%	14%	5%	57.5%	41%
Zero or Negative	less than 5%	12%	0%	2.5%	9%
SCORE STATISTICS					
Mean		28.79	15.28	29.13	16.71
	(indicated)[a]	(25.78)	(13.75)	(25.22)	(13.75)
Standard Deviation		5.13	3.76	5.23	2.99
	(indicated)[b]	(5.08)	(2.75)	(4.92)	(2.75)
Reliability	more than .70	.784	.798	.803	.690
Probable Error	no standard	1.60	1.14	1.57	1.12
SPEEDEDNESS					
% of Papers Completed	more than 90%	94%	100%	96%	100%

a - Midpoint of range between highest possible and expected chance score.

b - One-sixth of range between highest possible and expected chance score.

RESULTS

The results obtained from Experiment I are presented in summary form in Table IV. In addition to these numerical data the following impressions were obtained. Although impressions can not rightfully be considered to be as rigorous as data from tests of known reliability, they represent information that gives insights into the area under consideration.

1. The students viewing the films found it difficult to take adequate lecture notes.
2. The teacher who taught the Film I treatment sections did not spontaneously use the accepted methods of introduction and summary with the films.

The results from Experiment II are presented in Table V and from Experiment III in Table VI. The following impressions were gathered during the course of these experiments.

1. Other than carefully pre-viewing the film, little effort is required to properly instruct an undergraduate teaching assistant to conduct any particular class.
2. When instructed, the teaching assistant appeared competent to introduce and summarize a film presentation.

3. The notes used by the film teacher for his lecture, when duplicated and distributed, appeared to be accepted and used by the students as a record of the lecture in the same fashion as personally taken lecture notes.

The total sample used in this study is compared to the published (15) norm groups of enrolled students in Public and Private colleges of Engineering in Table IX. The figures in the norm groups are for students who have had no college experience, and the figures for this sample are for students who are enrolled in Engineering Problems, a course that is normally given to first term freshmen. Many of the students who are enrolled in Engineering Problems have had one or two terms of college work to make up deficiencies in entrance requirements before enrolling in this course.

The coorelation coeficients between the stratifying variable, Pre-Engineering Aptitude Test scores, and the two criterion measures, the Unit Test and Final Examination scores, are presented in Table X.

TABLE IV The results of Experiment I

1. Unit examination measuring the results of learning test number (580921)

   a. Summary table

Source	d.f.	s.s.	m.s.
Treatments	2	91.98	45.99
Levels	2	784.05	359.02
T X L	4	60.58	15.14
Cells	8	900.16	112.52
Within	87	1817.39	20.89
Total	95	2718.00	

   b. "F" tests

Variance Ratio $\frac{m.s._{Tr}}{m.s._{Within}} = \frac{45.99}{20.89} = 2.20$

F at 5% with 2/87 d.f. = 3.10

Variance Ratio $\frac{m.s._{L}}{m.s._{Within}} = \frac{359.02}{20.89} = 17.19$

F at 5% with 2/87 d.f. = 3.10

Variance Ratio $\frac{m.s._{TXL}}{m.s._{Within}} = \frac{15.14}{20.89} = 0.72$

F at 5% with 4/87 c.f. = 2.47

c. Table of cell, treatment, and level means, giving cell sample sizes and the amount of "lost" data.

			Treatment Groups			Levels
			Control	Film I	Film II	
Engineering Ability Levels	High	$\bar{\bar{x}}$	33.79	30.93	31.80	32.27
		n	14	14	5	33
		lost	1	0	0	1
	Medium	$\bar{\bar{x}}$	28.36	28.64	26.00	28.12
		n	14	14	5	33
		lost	0	0	1	1
	Low	$\bar{\bar{x}}$	27.21	24.14	25.20	25.61
		n	14	14	5	33
		lost	1	0	0	1
Treatment		$\bar{\bar{x}}$	29.79	27.90	27.67	
		n	42	42	15	
		lost	0	2	1	

2. Final Examination measuring retention test number (590202)

a. Summary Table

Source	d.f.	s.s.	m.s.
Treatments	2	21.80	10.90
Levels	2	291.69	145.84
TXL	4	313.86	78.46
Cells	8	555.35	69.41
Within	82	866.83	10.57
Total	90	1422.18	

b. "F" tests

$$\text{Variance Ratio} = \frac{\text{m.s.}_{\text{Tr}}}{\text{m.s.}_{\text{Within}}} = \frac{10.90}{10.57} = 1.03$$

F at 5% with 2/82 d.f. = 3.11

$$\text{Variance Ratio} = \frac{\text{m.s.}_{\text{L}}}{\text{m.s.}_{\text{Within}}} = \frac{145.84}{10.57} = 13.80$$

F at 5% with 2/82 d.f. = 3.11

$$\text{Variance Ratio} = \frac{\text{m.s.}_{\text{TXL}}}{\text{m.s.}_{\text{Within}}} = \frac{78.46}{10.57} = 7.42$$

F at 5% with 4/82 d.f. = 2.48

c. Table of cell, treatment, and level means, giving cell sample sizes and the amount of "lost" data.

			Treatment Groups			
			Control	Film I	Film II	Levels
Engineering Ability Levels	High	$\bar{\bar{x}}$	19.00	16.79	17.40	17.82
		n	14	14	5	33
		lost	1	1	0	2
	Medium	$\bar{\bar{x}}$	14.14	16.79	9.60	14.58
		n	14	14	5	33
		lost	0	1	1	2
	Low	$\bar{\bar{x}}$	14.00	13.00	16.00	13.88
		n	14	14	5	33
		lost	2	0	2	4
Treatment		$\bar{\bar{x}}$	15.71	15.52	14.33	
		n	42	42	15	
		lost	3	2	3	

TABLE V: The results of Experiment II

1. Unit Examination measuring the results of learning test number (590310)

a. Summary Table

Source	d.f.	s.s.	m.s.
Treatments	1	.03	.03
Levels	2	326.28	163.14
T X L	2	1.33	.66
Cells	5	327.64	65.53
Within	33	367.15	11.13
Total	38	694.79	

b. "F" tests

Variance Ratio $\frac{m.s._{Tr}}{m.s._{Within}} = \frac{.03}{11.13} = .00$

F at 5% with 1/33 d.f. = 4.14

Variance Ratio $\frac{m.s._{L}}{m.s._{Within}} = \frac{163.14}{11.13} = 14.66$

F at 5% with 2/33 d.f. = 3.29

Variance Ratio $\frac{m.s._{TXL}}{m.s._{Within}} = \frac{.66}{11.13} = .06$

F at 5% with 2/33 d.f. = 3.29

c. Table of cell, treatment, and level means, giving cell sample sizes and the amount of "lost" data.

			Treatment Groups		
			Control	Film	Levels
Engineering Ability Levels	High	$\bar{\bar{x}}$	32.71	33.14	32.93
		n	7	7	14
		lost	0	0	0
	Medium	$\bar{\bar{x}}$	30.86	30.43	30.64
		n	7	7	14
		lost	1	0	1
	Low	$\bar{\bar{x}}$	26.29	26.14	26.21
		n	7	7	14
		lost	2	0	2
Treatment		$\bar{\bar{x}}$	29.95	29.90	
		n	21	21	
		lost	3	0	

2. Final Examination measuring retention test number (590603)

a. Summary Table

Source	d.f.	s.s.	m.s.
Treatments	1	2.38	2.38
Levels	2	4.33	2.16
T X L	2	10.63	5.32
Cells	5	17.34	3.47
Within	30	154.56	5.15
Total	35	171.90	

b. "F" tests

$$\text{Variance Ratio} = \frac{m.s._{Tr}}{m.s._{Within}} = \frac{2.38}{5.15} = .46$$

F at 5% with 1/30 d.f. = 4.17

$$\text{Variance Ratio} = \frac{m.s._{L}}{m.s._{Within}} = \frac{2.16}{5.15} = .42$$

F at 5% with 2/30 d.f. = 3.32

$$\text{Variance Ratio} = \frac{m.s._{T \times L}}{m.s._{Within}} = \frac{5.32}{5.15} = 1.03$$

F at 5% with 2/30 d.f. = 3.32

c. Table of cell, treatment, and level means, giving cell sample sizes and the amount of "lost" data

			Treatment Groups		Level
			Control	Film	
Engineering Ability Levels	High	$\bar{\bar{x}}$	17.57	18.43	18.00
		n	7	7	14
		lost	0	0	0
	Medium	$\bar{\bar{x}}$	18.00	16.43	17.21
		n	7	7	14
		lost	1	0	1
	Low	$\bar{\bar{x}}$	18.00	17.29	17.64
		n	7	7	14
		lost	3	2	5
Treatment		$\bar{\bar{x}}$	17.86	17.38	
		n	21	21	
		lost	4	2	

TABLE VI The results of Experiment III

1. Unit Examination measuring the results of learning test number (590310)

a. Summary Table

Source	d.f.	s.s.	m.s.
Treatments	1	15.04	15.04
Levels	2	226.32	113.16
T X L	2	26.35	13.18
Cells	5	267.71	53.54
Within	18	243.25	13.51
Total	23	510.96	

b. "F" tests

Variance Ratio $\frac{m.s._{Tr}}{m.s._{Within}} = \frac{15.04}{13.51} = 1.11$

F at 5% with 1/18 d.f. = 4.41

Variance Ratio $\frac{m.s._{L}}{m.s._{Within}} = \frac{113.6}{13.51} = 8.38$

F at 5% with 2/18 d.f. = 3.55

Variance Ratio $\frac{m.s._{T\ X\ L}}{m.s._{Within}} = \frac{13.18}{13.51} = .98$

F at 5% with 2/18 d.f. = 3.55

c. Table of cell, treatment, and level means, giving cell sample sizes and the amount of "lost" data.

			Treatment Groups		
			Control	Film	Levels
Engineering Ability Levels	High	$\bar{\bar{x}}$	30.50	29.25	29.88
		n	4	4	8
		lost	0	0	0
	Medium	$\bar{x}$	29.25	31.50	30.38
		n	4	4	8
		lost	0	0	0
	Low	$\bar{\bar{x}}$	21.75	25.50	23.62
		n	4	4	8
		lost	0	0	0
Treatment		$\bar{\bar{x}}$	27.17	28.75	
		n	12	12	
		lost	0	0	

2. Final Examination measuring retention test number (590603)

a. Summary Table

Source	d.f.	s.s.	m.s.
Treatments	1	9.37	9.37
Levels	2	39.08	19.54
T X L	2	124.76	62.38
Cells	5	173.21	34.64
Within	13	145.75	11.21
Total	18	318.96	

b. "F" tests

Variance Ratio $\frac{m.s._{Tr}}{m.s._{Within}} = \frac{9.37}{11.21} = .84$

F at 5% with 1/13 d.f. = 4.67

Variance Ratio $\frac{m.s._{L}}{m.s._{Within}} = \frac{19.54}{11.21} = 1.74$

F at 5% with 2/13 d.f. = 3.80

Variance Ratio $\frac{m.s._{T \times L}}{m.s._{Within}} = \frac{62.38}{11.21} = 5.56$

F at 5% with 2/13 d.f. = 3.80

c. Table of cell, treatment, and level means, giving cell sample sizes and the amount of "lost" data

			Treatment Groups		Levels
			Control	Film	
Engineering Ability Levels	High	$\bar{x}$	18.75	15.75	17.25
		n	4	4	8
		lost	0	0	0
	Medium	$\bar{x}$	13.25	18.25	15.75
		n	4	4	8
		lost	1	1	2
	Low	$\bar{x}$	17.00	11.25	14.12
		n	4	4	8
		lost	2	1	3
Treatment		$\bar{x}$	16.33	15.08	
		n	12	12	
		lost	3	2	

TABLE IX Comparison of the total sample to norm groups

Percent Scoring Lower Than Selected Scaled Scores

Scaled Scores Pre-Engineering Ability Test (No. Right)	Public College Freshmen (N=5144)	Private College Freshmen (N=2043)	Total Public and Private (N=7187)	Sample for this study (N=175)
81	99	99	99	100
70	98	94	98	95
59	94	81	90	73
56	92	76	88	67
54	90	71	85	62
52	88	67	82	54
50	84	62	78	46
48	81	56	74	39
45	77	50	69	33
43	72	43	64	37
41	68	37	59	21
39	63	30	54	17
36	57	25	49	9
34	52	21	43	7
32	46	18	38	5
30	40	14	33	5
27	35	11	29	3
25	30	9	24	2
23	26	7	20	0
21	21	5	17	0
18	17	4	14	0
16	13	2	11	0
14	11	2	9	0
3	2	0	2	0
0	0	0	0	0

TABLE X Coorelation Coeficients Between Pre-Engineering Aptitude Test and the Unit Test and Final Examination Raw Scores.

	Unit Test	Final Exam	No. of Pairs
Experiment I	.598	.479	99
Experiment II	.677	.090	42
Experiment III	.501	.467	24

DISCUSSION

Experiment I

The results of Experiment I, given in Table IV, may be summarized as follows. It appears that: a) a generally experienced instructor, having never taught the course before, can achieve the same average results using minimum cost films as an experienced instructor, teaching in person, who has taught the course many times; b) there is no apparent benefit from showing the films a second time; c) the results are similar for the unit examination, which measures the effect of learning, and the final examination, which measures retention.

The same average results can, of course, be achieved many ways; for example, the high ability students of treatment A may score disproportionately higher than the high ability students of treatments B and C, which might be balanced by the low ability students of treatments B and C scoring higher than the low students of treatment A. This would result in a selectivity between the various treatments and the ability levels which appears in the analysis as interaction. No interaction was found in the unit examination data, and it will be argued that the observed interaction of the final examination data is extrinsic, i.e., not an in fact characteristic of the parent population.

The average criterion scores on both the unit and final examination obtained by the students at any level differ from the scores obtained by the students at any other level. Comparison of the mean scores of all the students at any one level with the mean scores of the students at other levels shows that the students who scored higher on the Pre-Engineering Ability test scored higher on both the unit and final criterion tests. This demonstrates that Pre-Engineering Ability test scores are indicative of performance, and score on this test is a suitable basis for stratification of the sample population in this experiment. The coorelation coeficients between the stratifying variable, Pre-Engineering Aptitude Test scores, and the criterion measures, the Unit Test and Final Examination scores, are shown in Table X. The instability and value of these estimates, from experiment to experiment and from test to test, should be noted. The size of these coeficients casts some doubt on the meaning of the interaction term because the predictor used to stratify the sample is not highly coorelated, in a stable manner, with the criterion tests. The fact that the oceficients are not higher also reduces the precision of the "F" test of the ratio of the mean square for treatments to the mean square for error. The instability of these coeficients, from experiment to experiment, is reflected by the erratic pattern of the values of the ratio of the mean square for levels to the mean square for error.

The above inferences are drawn from the "F" tests of Table IV, which show the ratio of the mean square for treatments to the mean square for error has a value which is less than reasonably attributable to chance, i.e., this ratio has a value less than that value which has a 5% probability of occurence by chance. The ratio of the mean square for levels to the mean square for error is greater than reasonably attributable to chance, i.e., this ratio has a value greater than that value which has a 5% probability of occurence by chance. These effects are identical on the unit examination and the final examination. The mean square ratios that test for interaction exhibit values less than reasonably (5%) attributable to chance for the unit examination but greater than reasonably (5%) attributable to chance for the final examination sample data. This apparent conflict between the results of the tests for the presence of interaction is not serious because of the large amount of data "lost" from the final examination data. The unit examination had 3 cases missing, in a total 99 cases, no 2 cases from the same cell, which would only slightly disturb the probability of the decision tests. The probabilities of the decision tests from the final examination are more seriously disturbed since the pattern and quantity of "lost" data was much more serious. Because the missing data does disturb the probability of the decision tests, the final examination results are not as valid as those of unit examination.

Tests that show there is no statistically significant difference between the sample data of the treatments may, by the rules of statistical inference, be extended to the treatment populations, thence to the parent population itself because the design assumptions are presumed to be met (see Method). The same is true of the tests indicating that statistically significant differences do exist between the levels. The processes of extending inferences from sample data to whole populations is not as simple when dealing with interaction. Although results showing that interaction does not exist can be immediately extended from sample to population as before, results showing that interaction does exist need careful interpretation. Significant interaction in the sample data ("observed interaction") may be intrinsic, i.e., a true characteristic of the treatment populations, or it may be extrinsic, i.e., due in whole or in part to extraneous factors that systematically effect certain portions of the data.

The latter seems to be true of the observed interaction of the final examination data. The "lost" data, caused by students dropping the course between the time of the unit examination and the final examination, is predominately in the Low ability level. The six students who dropped the course between the time of the unit examination and the final examination averaged 5.46 raw score points on the unit examination below the mean of the cell in which their "ability" placed them, and since the average difference between ability levels on this examination is 3.32 raw score points, it is seen that the data "lost" on the final examination is data from persons who scored very low compared to the rest of their own cell.

This pattern of data suggests dropping was done by the lower of the Low ability level students because they feared failure. This loss of data systematically affects these cell mean scores because usual practice is to substitute the cell mean for each piece of lost data and reduce by one the degrees-of-freedom of the error term. Removal of the potentially lowest scores from any cell and replacing them by the cell average score, computed on the basis of the remaining scores, gives the cell a systematic upward bias. This effect on the Low ability level acts between the time the students become aware they are failing (after the first, or unit, examination) and the final examination. The final examination data are, therefore, the most affected.

Fundamentally, interaction is caused by non-uniformity of the differences between the cell means throughout the table; the upward bias of certain cells causes observed interaction. Table IV, 2, c, of final examination data shows that the low ability cell of the Film II treatment, for example, not only exhibits a mean well above the average of the cell means of the next higher ability level, but that this cell "lost" forty per cent of the cases that were available earlier. This table also shows the low ability cells lost twice as much data as did the medium or high ability cells at the time of the final examination, although the unit examination data show all levels lost exactly the same amount of data. Not only is a mild systematic upward bias acting on the lower strata, a definite and strong upward bias is acting on one particular portion of the Table. It then appears that the observed interaction of the final examination data is at least in part extrinsic, which is to say that the observed interaction between treatments and levels of the final examination sample data is not a true characteristic of the treatment populations.

Certain impressions were formed from observing students in both film treatment and non-film treatment sections. In the film treatment sections many of the students attempted to take lecture notes during the films. These notes appeared both illegible and incomplete due to the semi-darkened room and the rapid development of the material on the film. It was felt that duplicated lecture notes would have overcome this difficulty.

The instructor who taught the film treatment sections was generally experienced though he had never taught this particular course before. He was given the films to use in lieu of the lectures, which he previewed and then presented to his classes. He did not use the methods of introduction, discussion, and summary which have come to be considered standard practice for this type of audio-visual material. (10)

The two characteristics of the teaching method of the film treatment, inability of the students to take satisfactory notes, consequently having no materials for review, and the lack of an introduction and summary to the films, are generally considered undesirable. (10, 11) If the lack of these characteristics in the teaching method do, in fact, hinder learning and in consequence depress criterion scores, the average scores of the film treatment groups could be improved by overcoming these detriments. In Experiment II and Experiment III the teachers of the film treatment sections gave careful introductions and summaries to complement the film presentations, and distributed to the students duplicated notes on the lectures.

Experiment II

The numerical results of Experiment II were much the same as Experiment I. The interpretation of the results of Experiment II may be summarized as follows: a) a carefully supervised undergraduate teaching assistant, using films and duplicated lecture notes, can produce the same effect as an experienced teacher giving the same lecture in person, b) this result is similar for the unit examination, which measures the effect of learning, and the final examination, which measures retention.

The above inferences were drawn from the "F" tests of Table V 1, b and V 2, b, which show the ratios of the mean square for treatments to the mean square for error has values which, for both the effect of learning and the effect of retention, are less than the maximum value reasonably (5%) attributable to chance.

The results of decision tests from the unit examination data are tabulated in Table V 1, b, and correspond very closely to Table IV 1, b, which presents similar results for Experiment I. These results show that the ratio of the mean square for levels to the mean square for error is larger than the maximum value reasonably (5%) attributable to chance, which indicates that the score on the test used to stratify the students into levels (Engineering Aptitude Examination Form ZPA) is a suitable basis for this purpose. The "F" test for interaction, the ratio of the mean square for interaction to the mean square for error, yields a result that is less than the maximum value reasonably (5%) attributable to chance indicating that no significant interaction is present.

The Final Examination data do not present a clear picture because approximately 15% of the data have been "lost." It is unrealistic to base decisions on results that have been so attenuated. This final examination data show that the ratio of the mean square for treatments to the mean square for error was smaller than the maximum value reasonably (5%) attributable to chance, as was the ratio of the mean square for levels to the mean square for error, and the ratio of the mean square for treatments by levels to the mean square for error. The values of these ratios indicate that there was no significant difference between the score obtained by the two treatment groups, there was no significant difference between the scores of the students at the different levels, and there was no significant interaction.

The first of these final examination results, which indicates no significant difference between the retention of the two treatment groups, is exactly the same result as obtained in Experiment I. The lack of significant difference between the scores of the different levels, and the lack of significant interaction is not the same result as was obtained in Experiment I. It appears that the reason for this discrepency is the large amount of lost data, and the cell location from which the loss took place. In this experiment the pattern of the "lost" data generally conformed to the pattern of Experiment I, except it was uniformly lost from the low and intermediate ability cells, rather than being concentrated in a particular cell. Apparently, the upward bias on the lower and the intermediate ability level cells of the table acted in such a uniform manner such that interaction is not introduced. The extremely low value of the coorelation coeficient between the stratifying variable and the final examination score tends to explain the lack of a significant "F" when testing levels effect. This coorelation coeficient would also cast doubt on the meaning of the interaction effect, had it been present.

Two impressions were obtained during the course of the experiment. First, that little effort, in addition to carefully previewing the film with the teaching assistant, is necessary to instruct properly the teaching assistant to conduct any particular class. Second, that the teaching assistant appeared competent to introduce and summarize film presentations. These impressions suggest that it would be possible for one experienced teacher to utilize and control a number of teaching assistants. The fact that duplicated lecture notes were accepted and used by the students in the same manner as personally taken notes indicates that the previously noted difficulty which the students have in taking notes from film, can be satisfactorily overcome.

It is interesting to note in passing that the film treatment sections of Experiment I were unable to take satisfactory lecture notes from the films, and they were neither supplied with lecture notes nor given an introduction and summary to the film; yet compared to their own control group they performed as well as did the film treatment students in Experiment II compared to Experiment II control group. This suggests that notes and the introductions and summaries to films may not be as important with this type and use of film as these same pedigogical techniques used with an individual film injected in a sequence of study to develop a particular concept. This suggestion warrants detailed study.

Experiment III

The results of the "F" tests of Experiment III may be summarized as follows: a) an instructor using the films, with duplicated notes and adequate introductions to and summaries of the films, can achieve the same results as an experienced senior staff member who has taught the course many times before and who uses his own teaching methods, b)these results are similar for the unit examination, which measures the effect of learning, and the final examination, which measures retention.

The above inferences were drawn from the ratios of the mean square for treatments to the mean square for error taken from the unit examination and the final examination data in Table VI, 1, b and VI, 2, b. Both these ratios had values less than the maximum values reasonably (5%) attributable to chance. Further analysis of Table VI, 1, b indicate that the ratio of the mean square for levels to mean square for error is larger than the maximum value reasonably (5%) attributable to chance, so, like the unit examination data from Experiments I and II, the test scores used to stratify the sample into ability levels appear suitable for this purpose. The unit examination data also show that the ratio of the mean square for interaction to the mean square for error is smaller than the maximum value reasonably (5%) attributable to chance, so no significant interaction is present in this data.

The final examination data are presented in Table VI, 2, a, b, and c. It should be noted immediately that there is a very high percentage of "lost" data, over 20%. With such severely attenuated data, and only four cases in each cell to start with, the results of these calculations are useless for decision purposes. However, it is interesting to note the results are similar to Experiment I and II in that there is no significant difference between treatments, and similar to Experiment I in that there is observed interaction between treatments and levels. In the Experiment III final examination data, however, there is no significant difference between levels.

The ratio of the mean square for treatments to the mean square for error from the final examination data is less than the maximum value reasonably (5%) attributable to chance, which indicates that there was no significant difference between the performance of the two groups with respect to retention. The ratio of the mean square for levels to the mean square for error from the final examination data was less than the maximum value reasonably (5%) attributable to chance, which indicates that there was no significant difference in performance between the various levels. As before, this fact reflects on the meaning of the interaction effect. The ratio of the mean square for interaction to the mean square for error was greater than the maximum value reasonably attributable to chance, which indicates observed interaction is present in the data taken from the scores on the final examination.

CONCLUSIONS

Two major conclusions seem to emerge clearly from the evidence that has been discussed.

1. Students can learn both conceptual and factual material from this "minimum cost" type of film, and they apparently learn as well as from the same lecture presented live under similar circumstances.
2. This type of film offers promise of being a) economically feasible on a self-amortizing basis for repetitive, multiple section courses b) helpful for teacher training, and c) useful for extending the influence of superior teachers.

FUTURE RESEARCH

The future research that needs to be done, to verify and establish the limits to the conclusions discussed here, is mostly extensive rather than intensive. The implications of this research, if they may be extended to whole courses, are sweeping and important. However, before any extension can be made, much additional experimental work needs to be done. The work reported here is in the nature of a pilot study--one study unit of one course.

The sort of extensive topics that need examination are as follows: Will students respond throughout whole courses in the same manner they have responded in the pilot study? Will they respond in other types of courses, i.e., lecture-recitation courses of a mathematical nature and lecture-recitation courses of a verbal nature, in the same manner as they have in this computation laboratory course? What are the attitudes of teachers using this teaching method and students studying by this method? What is the effect on drop-outs? The sales and rental revenue situation for films that are produced needs to be examined. The acceptability of these films in other institutions and the resulting inter-institutional relationships need examination. The usefulness of films of this type as a stabilizing influence and as a core around which teacher training may be organized for beginning teachers needs examination.

Of these many topics for research, two distinct research problems seem to stand above the others. First, the films for the entire course should be completed and the research described here should be substantially repeated for the whole course of study; second, because this is a computation laboratory course, where the content material is introduced by lecture, duplicated notes, problems, and individual help, it should be determined which of these forms of teaching carries what portion of the burden of the instruction. We should attempt to evaluate how necessary are the notes, introductions, and summaries to the films. We may find, for example, students would succeed equally well on criterion tests if only given lectures; perhaps just being given the problems to do would be sufficient.

With these primary questions answered, further extensive study can then proceed in an orderly and fruitful manner.

APPENDIX A

THE COURSE

Catalogue Description

59:2 Engineering Problems 2 S.H.

An elementary course stressing computation methods and procedures as applied to simple engineering problems. The use of logarithms and slide rule. Required of first year students.

Administrative Arrangements

The course meets two times a week for 16 weeks, three hours each meeting. At each meeting the students receive a lecture and problems which are to be completed correctly before the end of the class period.

Schedule:

1.	3 weeks	Logarithms, computational techniques and solution of triangles.
2.	4 weeks	Slide rule, study all conventional scales (approximately 13 scales).
3.	3 weeks	Slide rule, efficient use on problems.
4.	3 weeks	Slide rule, trigonometric problems.
5.	3 weeks	Review and examination.

Note: The above topics do not appear discretely in this order since the examinations, review, and considerations of efficient computation techniques are spaced throughout the term to maximize the learning.

Population:

The student population enrolled in this course at the State University of Iowa is of wide range in intelligence and aptitude. The population ranges between 75 and 125 students and is usually divided into 5 or 6 sections. The course is taught both semesters.

APPENDIX B
(Statistical Design Considerations)

Because of the nature of statistical inference the parent population and the random sample need to be carefully defined. Because the treatment groups of the "treatments by levels" experiments are constituted by the counting-off method, a hypothetical population must be created to fit the sample used for each experiment, these hypothetical populations must be compared to the real parent populations, and the assumptions of the experiments themselves must be considered individually. The arguments for the foregoing must necessarily be carried out at a descriptive level with descriptive statistics, because adequate statistical decision tests do not exist to make the necessary comparisons.

The"treatments by levels" design is made up of a number of treatment groups, each stratified into levels according to some measure taken on the population. The stratifying variable may be chosen on empirical or rational grounds, but the higher the correlation between the stratifying measure and the criterion measure, the finer will be the discriminations of the decision tests.

Each treatment group must have the same proportion of cases in each of the levels as does the parent population, and the cases used throughout the experiment at each level must be randomized between the treatment groups. When dealing with a population of unknown parameters, the cases are first ordered according to the stratifying variable and then "counted-off" into the treatment groups; then the treatments are assigned to these groups in a random manner. The levels are determined so that each treatment group has the same proportion of cases at every level. Instead of meeting the criteria that the treatment subgroups are independently randomized and are representative and proportional samples with respect to the parent population, these criteria are met with respect to some hypothetical population whose parameters are defined to fit the sample that has been drawn. The usefulness of the results then hinge on how closely the hypothetical population, defined for convenience, matches the true parent population about which it is desired to make inferences; because inferences from the experiment may be logically extended only to the hypothetical population--further extension to the parent population must be made on a judgment basis.

The other assumptions basic to this design are that the distribution of the criterion measure of the subgroup population be normal with an homogenous variance. The Norton study, as reported by Lindquist (12), shows that if all assumptions are met except homogeniety of form and variance, the resulting distribution is very close to an F-distribution.

This means that considerable latitude can be allowed in meeting these restrictions without perceptibly disturbing the probabilities of the decision tests.

The parent population, for purposes of our experiment, is all the students who have taken, and who will take, Engineering Problems at the State University of Iowa. The random sample is the students who take Engineering Problems during the 1958-59 school year. No factor is known to be biasing the sample in any systematic manner.

The course was taught during both semesters in five sections meeting at three separate hours. Two sections met at each of two meeting times, and the fifth section met on different days. At each of the hours that the two sections met, one was randomly designated experimental and the other designated control. All the students meeting at these times were pooled, and assigned to either the treatment or control section by the counting-off method on the basis of their ability level as judged by their scores on the Pre-Engineering Ability Test (Form ZPA). The counting-off method had to be used to establish the various treatment groups because the parameters of the parent population are not known.

The "levels" of the treatment groups were obtained by dividing the "matched" treatment groups into thirds. The method employed for constituting treatment groups and levels had the advantage of assuring both cell proportionality and representative experimental and control groups. However, this method cannot ensure that the single section can be matched with any control group, or the pairs of sections, one pair meeting in the morning and the other pair in the afternoon, can be brought together in a single experiment except through a covariance design. The covariance design is not desirable for this experiment because it does not permit examination of interaction and involves restrictive assumptions.

The portion of the sample available during the spring semester showed an unbalance of high ability students at certain meeting times; because of scheduling difficulties there was no way to balance the paired sections meeting at different hours. At any particular meeting time the two treatment groups generated by the counting-off method, together with subsequent division into thirds for levels, produced representative treatment and control sections for valid tests of hypotheses concerning a hypothetical population defined to fit the sample. The single section meeting on different days became unusable for purposes of either experiment or control because it could not be matched with another section. In Table VII the "best estimate" of the aptitude score of the different hypothetical populations are compared to the "best estimate" of the aptitude score of the parent population.

TABLE VII Comparison of the "best estimates" of the range and mean of the Pre-Engineering Aptitude scores of the Parent and Hypothetical Populations of the various experiments.

	mean	range	sample size
Parent Population	50.26	24.79	179
Hypothetical Population; Experiment I	47.15	24.70	99
Hypothetical Population; Experiment II	58.50	34.79	42
Hypothetical Population; Experiment III	49.88	26.74	24

Because the difference between the mean scores of all the hypothetical populations and the mean score of the parent population is small, it was felt that it was possible to extend the conclusions to the parent population on a judgment basis.

During the spring semester the "treatment by levels," or "factorial," design was used with both of the pairs of experimental and control sections that were created. The assumptions that must be made for this design, representative sampling and cell proportionality, are strictly met by the methods used to constitute the sections that correspond to the treatment groups. There are no good decision tests available with which to examine whether the data meet the additional assumptions of subgroup normality and equal subgroup variances. Considerable latitude is permitted in meeting these assumptions because of the findings of Norton cited earlier. (10) Inspection of the data reveals no marked discrepancies.

During the fall semester more uniformity occurred in the distribution students of various ability levels from section to section than during the spring semester. This presented the possibility of using the whole five sections for experimental purposes within one "treatments by levels" experiment. When the treatment groups were established, in the same manner that has already been described, the cutting scores between the levels and the mean scores of the cells at the same level fell very close together for all the groups, both experimental and control. And although adequate statistical decision tests are not available to compare morning and afternoon groups to ensure that the data may be pooled into one design, along with the single section, a strong case can be made from descriptive statistics. The comparison of cutting scores and cell means is shown in Table VIII.

The figures in Table VIII show conformity from treatment group to treatment group and from level to level. The widest range in cutting scores is two points and occurs when cutting between both the high and medium and the medium and low score groups. This value is reasonable when it is compared to the standard error of the test which is 3.9 raw score points. The range of cell mean ability scores is 3.12 units and occurs at the lowest level. The range of the treatment group mean ability scores is 2.56. On the basis of these descriptive statistics, it appears reasonable to accept that the treatments are representative samples from the hypothetical population defined to fit this total sample. The assumption of cell proportionality is strictly met by the method used to divide treatment groups. Only the "within cell skew" has not been dealt with, and the assumptions of this design call for only a normal distribution, with an homogeneous variance, of

TABLE VIII Comparison of Pre-Engineering Aptitude raw scores for the various experimental sections for the Fall Semester Sample.

Levels	Statistic	Control A.M.	Sections P.M.	Film A.M.	Sections P.M.	Single Section
High	$\bar{\bar{x}}$	59.57	59.86	60.29	58.00	61.80
	cutting score	53	52	53	51	53
Med.	$\bar{\bar{x}}$	46.29	47.14	46.43	46.14	47.20
	cutting score	42	41	42	40	42
Low	$\bar{\bar{x}}$	35.00	35.00	34.28	34.57	37.40
Total	$\bar{\bar{x}}$	46.95	47.33	46.90	46.24	48.80

the criterion measure of the population defined by the treatment subgroups, which the Norton study shows may be considerably relaxed without invalidating the tests. These data were examined subjectively for skew within the various cells and none was detected.

APPENDIX C

Computations for Economic Analysis

Present Method

The $6000 estimate shown in Table I as the yearly cost of teaching Engineering Problems was made by Professor J. W. Howe, Head, Department of Mechanics and Hydraulics, State University of Iowa, based on salary and time allotments during the 1956-57 school year. During this school year five sections were taught during both the fall and spring semesters. All sections had teachers of academic rank of instructor or above.

Under the present method of instruction, each meeting of each section is handled by a single experienced teacher. Each section meets twice a week for three hours each meeting, therefore, six clock hours per week are required for direct teaching duties. On the average, about two hours per week per section are required outside of class for planning, developing assignments and tests, grading examinations, and conducting student conferences. The time required for five sections taught one academic year of two semesters amounts to a total of about 1280 hours. This 1280 hours of work is all done by professional level people, although it is not all professional level work.

Proposed Method

The teaching costs shown in Table I for the Proposed Method are developed in the following manner. An assistant will have to devote about eight hours per week on the average to this work for thirty-two weeks to include both semesters of the academic year. The eight hours may be broken down. There will be a one-half hour orientation before each of the three hour meetings which occur twice a week; this totals seven hours per week in orientation and direct teaching work. The other indirect duties involved in assisting in this course, such as giving and reading examinations, developing problems, etc., should average about one additional hour per week.

At present, at this institution, the top rate of pay for student assistants who grade papers and read examinations is $1.25 per hour. A higher salary would be appropriate for teaching assistants who, in addition to grading papers, operate projectors, monitor classrooms, and give some individual help to students during the class period. A fair rate of pay to use for estimating purposes would be from $1.50 to a top of $2.00 per hour. This works out to be $384.00 to $510.00 a year, which, considering the nature of our estimates, should be rounded off to $400.00 to $500.00 per year. It would make no difference to the estimate whether the teaching assistant were hired on a contractual basis or on an hourly basis. The additional prestige of an appointment may be desirable to attract the best students to apply for these positions.

Supervision of the course by a ranking faculty member is estimated to require approximately one-fourth of one man's time for which $1600.00 yearly seems a reasonable estimate.

The time estimate for the proposed method has two categories, the work of the assistants, the nature of which must be all subprofessional, and the work of the supervising teacher, which should be arranged so that it contains as little subprofessional work as possible. The comparison of interest is the professional time required by each of the methods. The subprofessional time requirements of the proposed methods have already been detailed when the monetary estimates were made. This estimate may be extended to obtain a total of about 1280 hours for the subprofessional work required.

About 10 hours of work per week will be required from the supervising teacher to handle five sections. This time may be broken down as follows: one-half hour per class meeting will be needed for orientation of assistants; one-half hour per class meeting will be devoted to the professional development of the assistants; one hour per class meeting will be needed to prepare the material for the orientation and professional development meeting. Each time a class meets the supervising teacher should spend ten to fifteen minutes visiting the class to support and aid the teaching assistant with his day to day work; one hour per week will be required for conferences with individual students and assistants; and an average of one hour per week will need to be devoted to producing new films. For five sections for an academic year of two semesters, the total time spent by the supervising teacher would be about 320 hours.

APPENDIX D

Film Production

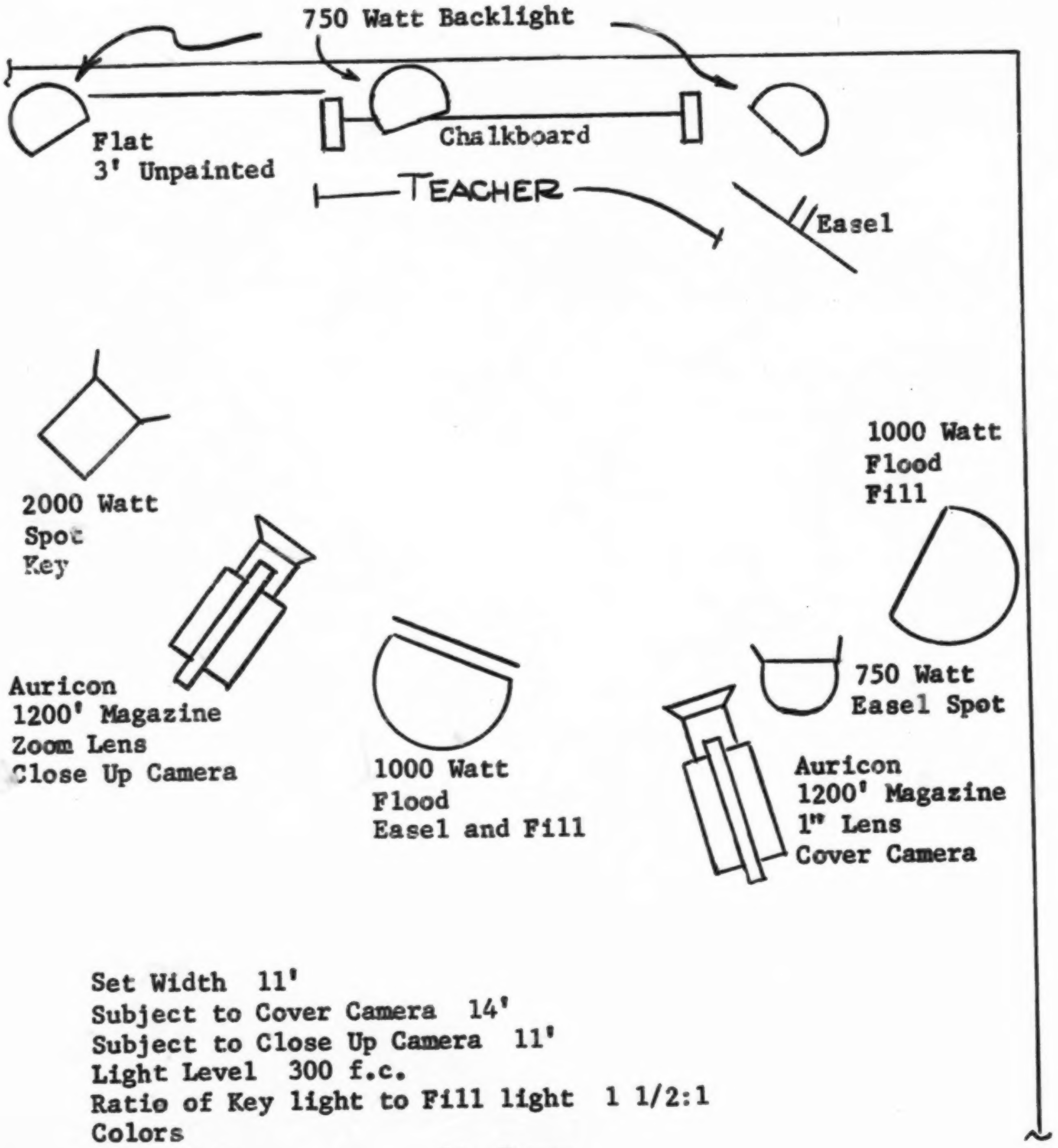
Six, sixteen millimeter synchronous sound motion pictures were produced for this study. They averaged about twenty-five minutes running time, the longest thirty minutes, the shortest twenty minutes. They were all made with a two camera technique, one camera equipped with a zoomar lens. A diagram of the set is shown in Figure 5. The parameters of interest to cameramen are included on the set diagram.

The films were produced on the sound stage in the basement of East Hall on the campus of the State University of Iowa. Because of the limitation of the stage, several compromises had to be made:

1. It was necessary to "pan" with the "cover camera" because the sound stage was not large enough so that the "cover camera" could include the entire set.
2. The blackboard was not large enough to work completely through some of the material to be displayed. This was overcome by repeated erasing and by using display cards on an easel.
3. The sound stage is not large enough to locate the close-up camera (equipped with a zoomar lens) on the right hand side of the set. A view from the right of

of a right handed teacher while writing on the blackboard would have made the material more easily visible.

Production of the films was attempted in one scene to eliminate special transitional shots necessary for editing, thus reducing editing time and saving film. Experience with the films indicates that this technique is possible if the scene length is under about twenty minutes. If the scene length becomes longer the discomfort from the heat and perspiration becomes so severe, from the 300 foot candles of illumination, that it is practically impossible for the "teacher" to give a smooth performance. This would not be crucial except that the last few minutes of the film are usually reserved for summary, and, it seems desirable to have this portion of the presentation smooth and forceful for lasting effect.


750 Watt Backlight
Flat
3' Unpainted
Chalkboard
TEACHER
Easel
2000 Watt
Spot
Key
1000 Watt
Flood
Fill
Auricon
1200' Magazine
Zoom Lens
Close Up Camera
1000 Watt
Flood
Easel and Fill
750 Watt
Easel Spot
Auricon
1200' Magazine
1" Lens
Cover Camera

Set Width 11'
Subject to Cover Camera 14'
Subject to Close Up Camera 11'
Light Level 300 f.c.
Ratio of Key light to Fill light 1 1/2:1
Colors

Wall	Lt. Brown
Flat	Off White
Chalkboard	Black
Teacher's Suit	Lt. Blue

Figure 5. Set Diagram

BIBLIOGRAPHY

1. Dale, E., Dunn, F. W., Hoban, C. F., Jr., Schneider, E., *Motion Pictures in Education*, American Council on Education, Committee on Motion Pictures in Education, H. W. Wilson, 1937.

2. Hoban, C. F., Jr., Van Ormer, E. B., "Instructional Film Research", *Technical Report No. SDC 269-7-19*, 1951.

3. George, C. S., Jr., Parden, R. J., "Management Training Films", unpublished dissertation, State University of Iowa, 1953.

4. Harshbarger, H. C., Becker, S. L., *Teaching by Discussion on Television*, duplicated, Television Center, State University of Iowa, 1955.

5. Kumata, Kidaya, *An Inventory of Instructional Television Research*, Educational Television and Radio Center, Ann Arbor, Michigan, 1956.

6. Seibert, W. F., "Closed-Circuit Television Instruction in Mechanical Engineering", *Journal of Engineering Education*, V48, N 7

7. Greenhill, L. P., Kepler, L. F., A Study of the Feasibility of Local Production of Minimum Cost Sound Motion Pictures, *SPECDEVCEN Project 20-E-4*, 1955

8. Anderson, G. R., Vander Meer, A.W., "A Comparative Study of the Effectiveness of Lessons on the Slide Rule Presented Via Television and In Person", *The Mathematics Teacher*, 1954, 47, 323-327

9. ________________, *CCTV Instruction*, preliminary reports duplicated, Purdue University, 1936.

10. Dale, E., *Audio-Visual Methods in Teaching*, Revised Edition, The Dryden Press, 1954.

11. Stroud, J. B., *Psychology in Education*, Longmans, Green and Co., 1956.

12. Lindquist, E. F., *Design and Analysis of Experiments in Psychology and Education*, Houghton Mifflin, 1953.

13. Hoel, P. G., *Introduction to Mathematical Statistics*, 2nd Edition, Wiley, 1954.

14. Snedecor, G. W., *Statistical Methods*, Iowa State College Press, 1946.

15. ______________, Manual for the Pre-Engineering Ability Test Form ZPA, Educational Testing Service, Princeton, N.J., 1952.